

Complexity and characteristic frequency studies in ECG signals of mice based on multiple scale factors

YANG XiaoDong^{1,2}, HE AiJun², LIU Peng¹, SUN TongFeng¹ & NING XinBao^{2*}

¹*School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China;*

²*School of Electronic Science and Engineering, Institute for Biomedical Electronic Engineering, Nanjing University, Nanjing 210093, China*

Received October 6, 2010; accepted February 10, 2011

Existing methods of physiological signal analysis based on nonlinear dynamic theories only examine the complexity difference of the signals under a single sampling frequency. We developed a technique to measure the multifractal characteristic parameter intimately associated with physiological activities through a frequency scale factor. This parameter is highly sensitive to physiological and pathological status. Mice received various drugs to imitate different physiological and pathological conditions, and the distributions of mass exponent spectrum curvature with scale factors from the electrocardiogram (ECG) signals of healthy and drug injected mice were determined. Next, we determined the characteristic frequency scope in which the signal was of the highest complexity and most sensitive to impaired cardiac function, and examined the relationships between heart rate, heartbeat dynamic complexity, and sensitive frequency scope of the ECG signal. We found that all animals exhibited a scale factor range in which the absolute magnitudes of ECG mass exponent spectrum curvature achieve the maximum, and this range (or frequency scope) is not changed with calculated data points or maximal coarse-grained scale factor. Further, the heart rate of mice was not necessarily associated with the nonlinear complexity of cardiac dynamics, but closely related to the most sensitive ECG frequency scope determined by characterization of this complex dynamic features for certain heartbeat conditions. Finally, we found that the health status of the hearts of mice was directly related to the heartbeat dynamic complexity, both of which were positively correlated within the scale factor around the extremum region of the multifractal parameter. With increasing heart rate, the sensitive frequency scope increased to a relatively high location. In conclusion, these data provide important theoretical and practical data for the early diagnosis of cardiac disorders.

ECG, multifractality, complexity, frequency scale factor, characteristic frequency

Citation: Yang X D, He A J, Liu P, *et al.* Complexity and characteristic frequency studies in ECG signals of mice based on multiple scale factors. *Sci China Life Sci.* 2011, 54: 544–552, doi: 10.1007/s11427-011-4173-y

Nonlinear analysis of biomedical signals is more commonly used than traditional linear analysis methods in biomedical research for analysis of essential physiological functions. Babloyantz *et al.* [1] first used several independent nonlinear dynamic measures to study four normal electrocardiograms qualitatively and quantitatively. In that study, analysis of phase space, Poincare section, correlation dimension, Lyapunov exponents, and Kolmogorov entropy demon-

strated that the normal human heart was not a perfect oscillator, but rather a chaotic system, and the electrical signals of the heart often developed into the shapes of the so-called strange attractors in reconstructed phase space as they evolved over time. Nonlinear dynamic parameters have been successfully applied to studies of cardiac electrical activity, which suggest that ECG signals exhibit at least partly chaotic characteristics [2–4]. Nonlinear dynamic parameters are also more effective than traditional linear approaches for discrimination of deterministic stochastic sig-

*Corresponding author (email: xbnings@nju.edu.cn)

nals. Nevertheless, existing data also suggest that a single parameter or fractal dimension may be insufficient to characterize such complicated behaviors as human bioelectrical activities. Because of their multiple dimensions, multifractal theories [5–16] are more effective for examining the diverse levels of the fractal system.

Amaral *et al.* [17] examined restrained neural function of a group of human subjects through drug experiments, and proposed that multifractal characteristics of human heartbeat R-R interval (HRV) were controlled by their autonomic neural function. These studies suggest that the multifractal phenomenon in heartbeat dynamics is an intrinsic feature of heart control mechanisms, and that these features are not changed with other factors such as external stimuli, physical activities, or other apparent alterations in behaviors including body position, dining, and variations of sleep states. During pharmacological restraint of human sympathetic nerves, the multifractal singularity spectrum of the patients ECG signal narrows, while its peak position is not altered compared with that of healthy subjects. During pharmacological restraint of human parasympathetic nerves, the multifractal singularity spectrum narrows significantly and its peak position moves right, suggesting a decrease in multifractality. Thus, the authors suggested that there are no definite mathematical equations that can describe complex dynamic behaviors in a physiological system that can generate multifractal signals. Further, even if these equations do exist, their solutions remain infeasible. Thus, experimental studies are required to develop improved modeling of the inherent mechanism of heartbeat dynamics to examine the nature of multifractal phenomena.

As an important extension method of multifractal theory, the multiscale approach was previously combined with multiscale entropy (MSE) by Costa *et al.* [18,19] for analysis of physiological time series. In that study, the authors examined the distributions of entropy measures with scale factors from HRV signals in healthy young and healthy elderly subjects, and found that at some scale factor or within a certain range, the nonlinear parameter resulted in better discrimination. Wang *et al.* [20] also proposed the multiscale multifractality (MSMF) method, and examined the distribu-

tion of singularity strength range with scale factor from ECG of 65 healthy humans. Our studies [21] pointed out that the essence of multiscale analysis is just to change the sampling frequency of the signals, and then explore them at diversified time levels.

In the present study, to examine the multi-layered cardiac electrical signals in the time domain in mice, we used multiscale analysis to change the sampling frequency (coarse-grained sampling frequency) of the signal through altering scale factors and then coarse-graining of the original time series. Furthermore, using pharmacological experiments in mice, we imitated different physiological and pathological states for detailed analysis of the distribution of ECG mass exponent spectrum curvature with scale factors under different sampling frequencies. Our results suggest that this method is effective for determining the scale factor and the sampling frequency, allowing determination of the characteristic frequency scope of ECG signals that best represent their complex fractal structures and dynamic features. In addition, the electrocardiosignals of the mice studied in this paper were high frequency electrocardiogram (HFECG) signals, as shown in Figure 1A, which contained frequency components higher than 100 Hz. The sampling frequency of the signals was set to 5 kHz due to higher heart rate of the mice compared with humans. Therefore, the frequency scope below 2500 Hz of the signals could be studied according to sampling theorem. In Figure 1B, we provided the probability density function (PDF) curve of this analyzed time series, whose mean value and standard deviation (SD) are -0.0532 mV and 0.0207 , respectively. Thus, our proposed methods allow us to determine valuable rules in heartbeat dynamics that may provide useful information on human disease states.

1 Theory

1.1 Multifractal theory

The multifractal method uses a spectral function to describe features of different levels of the fractal structure, and pro-

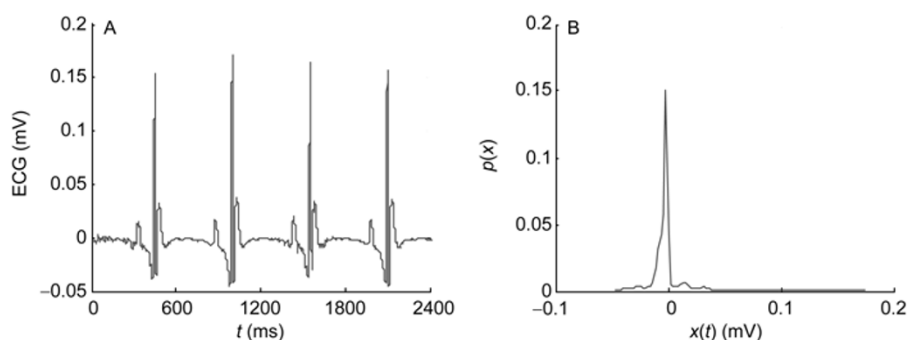


Figure 1 Time domain waveforms of the V4-lead high frequency electrocardiogram (HFECG) signal taken from a healthy mouse. A, The original time series. B, Probability density function (PDF) curve of the studied series displayed in A, which is exhibited as Gaussian distribution, approximately. The unit of the x -axis is millivolt (mV).

vides information on the global regularity based on localities. We use the standard box counting method proposed by Chhabra and Jensen [22] to estimate the multifractal spectrum. In brief, the measure chain is covered with segments of size L , and the probability in each of these segments is computed. The multifractal formula is then developed to describe the statistical properties of some measures in terms of their distribution of the singularity spectrum corresponding to its singularity strength.

The singularity measure (probability) of the segments whose measure value is T_i is defined as

$$P_i(L) = \frac{T_i}{\sum_{i=1}^N T_i} \sim L^{\alpha_i}, \quad (1)$$

where T_i is the i th share (box) of the time sequence whose measure chain is equally divided into N shares. In this paper, we took the mean value of voltage in each segment as the measure value of this segment. The exponent α_i is termed singularity strength, which reflects the singular degree of each small section in the fractal system.

Because of the non-uniformity of many physical fields, the strange attractor is also somewhat sparse and dense. In the evolution of the nonlinear process, subcluster courses of different orders of magnitudes often occur. This cluster process is a ubiquitous phenomenon in normal physiological activities [23]. Thus, chaos in different locations must be determined on diverse scales, i.e., we must consider the q th-order moment of the probability. To characterize the universality of the nonlinear structure, Renyi proposed the Renyi entropy [24]:

$$I_q(L) = \frac{1}{1-q} \log \sum_{i=1}^N P_i^q(L), \quad (2)$$

where L is defined by N^{-1} , which should converge to approximately zero for a large number of N , while q represents different scales. When $q \rightarrow 1$, Renyi entropy is simplified into Shannon entropy. Relative to the q factorial scaling exponent of probability measure, the generalized dimension D_q provides another description of singularity measure which is defined as

$$D_q = \lim_{L \rightarrow 0} \frac{-I_q(L)}{\log L} = \frac{1}{q-1} \lim_{L \rightarrow 0} \frac{\log \sum_{i=1}^N P_i^q(L)}{\log L}, \quad (3)$$

where D_q is also termed the multi-fractal dimension, which actually covers all dimensions involved in fractal theory, and further expands the meaning of fractal theory. In fact, D_q reflects the singular degree of spatial subsets by virtue of the q th powering and the summing operation to the singularity measure of these spatial subsets. Therefore, subsets with different scaling exponents can be distinguished through the changing of q values. The single fractal struc-

ture is divided into many regions with various singular degrees, which makes it more convenient to understand the inner structure of the fractal system from multiple levels. Thus, we introduced the scaling exponent of the moment, termed the mass exponent:

$$\tau(q) = \begin{cases} (q-1)D_q, & q \neq 1, \\ 0, & q = 1. \end{cases} \quad (4)$$

Since D_q and $f(\alpha)$ are the smooth functions of q and α , respectively, D_q is simply related to $f(\alpha)$ through the formula $\tau(q) = (q-1)D_q$ ($L \rightarrow 0$), which is determined by the Legendre transformation ($f(\alpha) = q\alpha - \tau(q)$ and $\alpha = d\tau(q)/dq$).

1.2 The mass exponent spectrum curvature

Variations of the exponent τ with different scales q form the mass exponent spectrum. Our previous studies [25] indicated that the curvature in mass exponents $\tau(q)$ is indicative of the degree of multifractality and nonlinearity. A nonlinear dependence of τ on the moments q is typical for nonlinear multifractal signals, while a linear dependence of τ on q signifies a linear fractal behavior. A straight line for τ - q corresponds to a monofractal signal. In this case, all the points in this curve belong to a same fractal dimension. This changing curvature indicates multifractality or fractal measures, which reflects the increment of complexity in fractal structure. The total bending degree of the curve can be regarded as the superposition from each part, which corresponds to different q values and fractal dimension. The larger the curvature, the stronger the nonlinearity of the original signal. As such, the curvature of the spectrum can be used as measure of nonlinear complexity for the original time series.

The estimation algorithm of the mass exponent spectrum curvature and its theoretical verifications were described [25]. In brief, according to the multi-fractal dimension theory described in Section 1.1, we produced the mass exponent $\tau(q)$ spectrum of the same ECG signal provided in Figure 1A, and the result is shown in Figure 2A. We divided the curve into the left and the right parts through $q=1$ ($q \in [q_{-\infty}, 1]$, $q \in [1, q_{+\infty}]$), where $q_{-\infty}$ and $q_{+\infty}$ were derived separately from the difference between each of their two successive function values in the curve that was less than a very small number (e.g., 1.0×10^{-6}). Next, the least squares fits of these data were determined on both sides. The two fitting lines intersect at point A. The angle formed is denoted as ϕ ($\phi \in (\pi/2, \pi]$), and the lengths of these two segments are l_1 and l_2 , respectively. The mass exponent spectrum curvature is then defined as

$$K_{\tau(q)} = \tan \phi. \quad (5)$$

For the above computations, to prevent the possibility

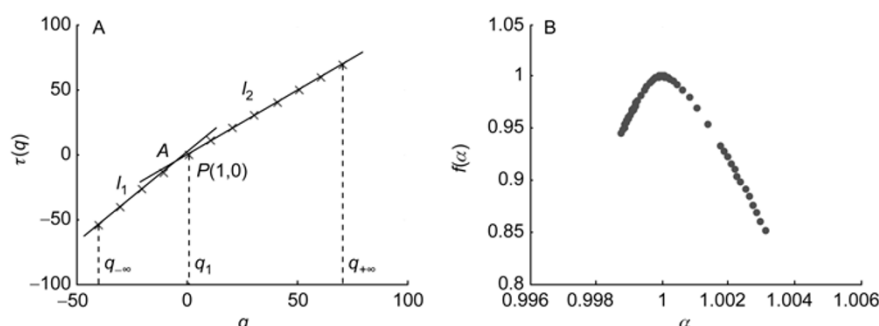


Figure 2 The multifractal curves of the V4-lead ECG signal taken from a healthy mouse. A, Mass exponent spectrum $\tau(q) \sim q$, made up of discrete data points marked with crosses, whose density is determined by the value selection of q . An apparent inflexion is formed at point $P(1,0)$ from which the curve is divided into two parts of the left and the right. The spectrum has different approximately linear behaviors for both negative and positive moments q , which extend to $q_{-\infty}$ and $q_{+\infty}$, respectively. B, Singularity spectrum $f(\alpha) \sim \alpha$, the α -axis width of which can be applied to calculate multifractal complexity.

that $P_i(L)$, which is given by eq. (1), equals zero, all the data in the sequence were uniformly preprocessed by adding the modulus of their minimum number before operations. Additionally, for reducing errors, in practice we estimate this multifractal spectrum $\tau(q)$ for a much denser set of values of q , even at each integral interval, whose density is much greater than suggested in Figure 2A.

The singularity spectrum $f(\alpha)$ versus its singularity strength α of this signal is provided (Figure 2B). The width of it is defined as $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$, where α_{\max} and α_{\min} are the maximum and the minimum singularity exponents, respectively. This indicator was used in previous studies to diagnose healthy and pathological heartbeat dynamics [17,26].

2 Experimental design

In this paper, $K_{\tau(q)}$ is used as the effective multifractal parameter to characterize cardiac electrical activities. We examined the distribution of this parameter with the frequency scale factor in mice that received pharmacological manipulation of cardiac states to determine the most sensitive nonlinear eigenvalues of ECG signals in discrimination of their complex dynamic features.

2.1 Subjects

Male mice from the Animal Center of Nanjing Medical University were housed under controlled ambient temperature ($(24 \pm 1)^\circ\text{C}$), with free access to food and water. The chemicals used were purchased from standard suppliers (Sigma Corporation). Mice ((20 ± 1.0) g) were divided into three groups. Animals in the various drug groups were injected two weeks prior to the tests. In Group A, mice ($n=18$) were injected with pure doxorubicin (DOX, 16.7 mg kg^{-1}), a highly effective drug used for treatment of tumors and hematological diseases, but which can produce dose-related

cardiotoxicity. In Group B, mice ($n=20$) were injected with Gal-PHEA-DOX couplings (GPD, 27.7 mg kg^{-1}), a drug used to reduce doxorubicin toxicity and maintain curative effects, but which can also decrease heart rate at certain doses. In Group C, the normal healthy control group (Con), mice ($n=12$) did not receive any drug or vehicle injection, and with sinus heart rhythm.

The main toxic side effect of doxorubicin is cardiac toxicity [27], which manifests clinically as acute arrhythmias and nonspecific electrocardiogram changes to decreased left ventricular ejection fraction. The acute cardiotoxicity occurs several hours after drug delivery, and presents as transitory alterations of cardiac electrophysiology and rhythm, including nonspecific changes of ST-T segment, lower QRS voltage, prolonged QT intervals, and arrhythmias in electrocardiogram waveforms [28]. In our experiments, we studied the period within two weeks after drug injection to assess acute side effects, during which symptoms of arrhythmia and abnormal electrocardiograph should occur. There may also be long-term effects of this drug such as pump failure and dilated cardiomyopathy following cumulative doses.

2.2 Experimental equipments and data processing

The ECG signals were recorded on an electrocardio-workstation NXD-2000 designed independently by the institute for Biomedical Electronic Engineering of Nanjing University [29]. This detection system has a broad frequency response and high sensitivity, and can detect and display frequency components higher than 100 Hz (100 Hz–5 kHz). The A/D conversion was 12 bits, and one datum was denoted by two bytes (the highest four bits were set to zero). Tests were performed in a quiet room with a constant temperature. Six electrodes were placed on each animal, three on the chest and three on limbs (the right leg was grounded), and were tightly fixed on the body to reduce myoelectric disturbances. The high frequency ECG signals were recorded,

amplified, and the analog signals transmitted to an A/D sampling card located in a PC workstation. Because of the higher heart rate of mice compared with humans, the sampling frequency of the card was set to 5 kHz per channel (lead). The sampling time lasted approximately 1 min, and the data were separated into 12 groups according to leads (approximately 300000 points per lead). The collected ECG data were not filtered before computations, for preserving their original characteristics. To eliminate the influence of initial unstable acquisition, we removed the first 2000 points of each data set and divided the remaining into N segments ($1/N \rightarrow 0$) with segment size $L=2$. The parameter q was set to vary from $-\infty$ to $+\infty$, and the steps changed according to the density of these points.

3 Results and discussion

3.1 Multiscale analysis

To determine the most appropriate sampling frequency, multiple scale factors were used to study the different time segments of the series. For the original time series $x = \{x_1, \dots, x_i, \dots, x_N\}$, we constructed the coarse-grained series of it, namely $\{y^{(\gamma)}\}$:

$$y_j^{(\gamma)} = \frac{1}{\gamma} \sum_{i=(j-1)\gamma+1}^{j\gamma} x_i, \quad 1 \leq j \leq N/\gamma, \quad (6)$$

and

$$F_c = \frac{f}{\gamma}, \quad F_s = \frac{F_c}{2}, \quad (7)$$

where γ is termed the scale factor. In the case of $\gamma=1$, the time series $y^{(1)}$ is equal to the original sequence x . In eq. (7), F_c and f are defined as the coarse-grained frequency and the sampling frequency, respectively, and the signal frequency F_s is one half of the value F_c .

In this experiment, we separately calculated three data lengths of the ECG time series, i.e., 277200, 201600, and 151200 data points. The sampling frequencies were all 5 kHz. Next, we computed the mass exponent spectrum curvature according to eqs. (5) and (6) for each of their coarse-grained data sets. First, we coarse-grained the ECG time series of healthy mice with a maximal scale factor 10 (Figure 3A). The parameter $K_{\tau(q)}$ for the three types of data number were not the same for each scale factor, but had their own spatial distribution along whose values tended to reduce with increasing γ . Next, we took the maximal scale factor 20 for computations. Irrespective of the data length, this case (Figure 3B) exhibited the same lowest section of the curves for $K_{\tau(q)}$ (extremum region) in which their absolute function values were maximum when the corresponding scale factors were in the range of $\gamma=7-12$. Subsequently, we continued to

increase the maximal scale factor to 30 and 50, respectively, and found that the y -axis values trended to increase with increasing γ , while the above-mentioned maximal scale factor range remained invariant. These two outcomes are demonstrated in Figure 3C and D, respectively. In GPD injected mice, the scale factor range in which the absolute magnitudes of $K_{\tau(q)}$ reached the maximum was in the range of $\gamma=10-15$ (Figure 3E), which was somewhat higher than that in healthy mice, and also remained unchanged. In DOX injected mice, the location of the scale factor in the extremum region was lower than both of the healthy and the GPD injected mice (Figure 3F), with a range within $\gamma=4-9$. In addition, as shown in Figure 3E and F, $K_{\tau(q)}$ was an oscillating function of scale factor γ . We consider that the origin of these oscillations may be cardiac injury resulting from drug administration which leads to unstable cardiac health status, which was most obvious in the DOX injected mice.

The above distributions of $K_{\tau(q)}$ to scale factors γ may indicate, as for the ECG signals of healthy mice, the most sensitive physiological information and complex fractal structures are likely to concentrate on the frequency scope $F_s=(357-208)$ Hz (based on eq. (7)), which is termed the characteristic frequency of the ECG. For the GPD injected mice that exhibited a heart rate lower than that of healthy animals, the most sensitive frequency scope (characteristic frequency) of the ECG was $F_s=(250-167)$ Hz. As heart rate decreases, the base-frequency and all its harmonic components of the ECG signals consequently decrease. At the same time the most sensitive frequency scope for characterization of complex dynamic features would also decrease to some extent. The opposite situation occurred in DOX injected mice that exhibited a higher heart rate, with a relevant frequency scope of $F_s=(625-278)$ Hz. We suggest that for multiscale analysis, construction of the coarse-grained series of original physiological time series, followed by examining the distributions of nonlinear parameters under different sampling frequencies, allows estimation of the most effective scale factor and frequency scope for expressing complex dynamic characteristics of the signal.

3.2 Experimental results

The multiscale multifractal analysis for the three comparative groups is shown in Figure 4. The maximal scale factor used was 50 in this computation, which corresponds to the signal frequency as low as 50 Hz according to eq. (7). Note that the differences of $K_{\tau(q)}$ values between each pair of the three classes of subjects reached the maximum within a certain scale factor range, i.e., $\gamma=7-15$, which corresponds to the frequency scope $F_s=(357-167)$ Hz for healthy and GPD injected mice in Figure 4A, and $\gamma=4-15$, which corresponds to the frequency scope $F_s=(625-167)$ Hz for GPD injected mice and DOX injected mice in Figure 4B. In these

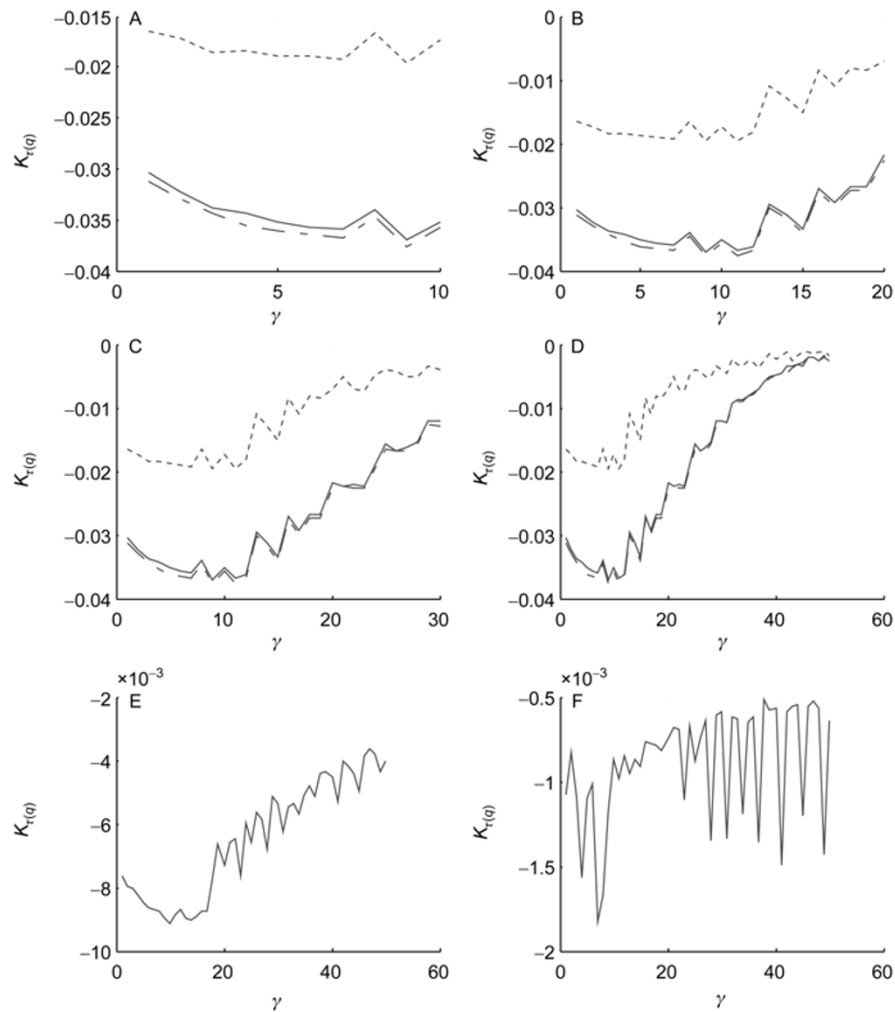


Figure 3 Curves of the mass exponent spectrum curvature $K_{\tau(q)}$ (mean value) of the V5-lead ECG signal versus scale factor γ . A–D, The healthy mice with maximal scale factors of 10, 20, 30, and 50, respectively. In each of these four plots, the three kinds of lines represent different data lengths. Solid line, 277200 data points; dash-dotted line, 201600 data points; dashed line, 151200 data points. For all the three types of data, the absolute magnitudes of the y-axis values achieve maximum when the corresponding scale factors are in the range of $\gamma=7-12$. The section's position of the lowest $K_{\tau(q)}$ value has no relationship with data length, but only with the scale factor γ . For each scale factor, the $K_{\tau(q)}$ value was the largest with the 151200 data points, and almost overlapped with the other two types of data points (smaller in the 201600 data points than in 277200 data points). We also found that the three curves representing different data lengths intersect at the scale factor between 50 and 53. E, The GPD injected mice for the 300000 data points of the ECG time series and with a maximal scale factor 50. The absolute y-axis values were maximal when the scale factors were in the range of $\gamma=10-15$. F, The DOX injected mice for the 300000 data points of the ECG time series and with a maximal scale factor 50. In this case, the maximal scale factor range is $\gamma=4-9$.

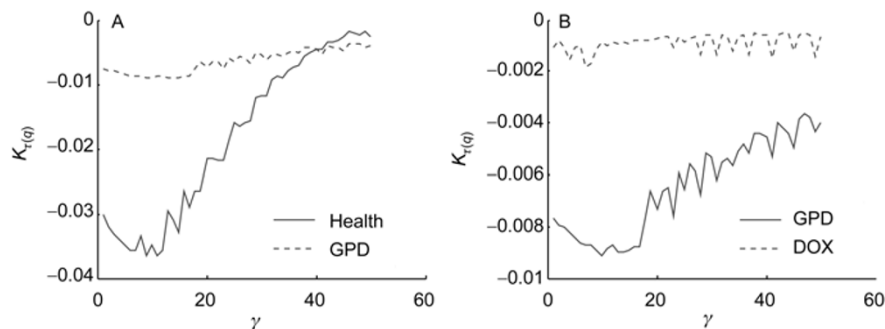


Figure 4 MSMF analysis of the ECG time series for each of the two compared pairs with maximal scale factor 50 and 300000 data point length (t -test, $P<0.05$). A, The healthy mice (solid line) and the GPD injected mice (dashed line). The two curves intersect at the scale factor between 40 and 42, from which we could not distinguish these two groups. B, The GPD injected mice (solid line, heart rate decrease) and the DOX injected mice (dashed line, heart rate increase). In both conditions in A, there is a respective scale factor range $\gamma=7-12$ for healthy mice and $\gamma=10-15$ for GPD injected mice in which the heartbeat dynamics is the strongest and also most sensitive to heart diseases. When the heart rate decreases (GPD injected mice), this scale factor range moves higher. While for the DOX injected mice (dashed line, B), the sensitive scale factor range is $\gamma=4-9$, which moves lower with increasing heart rate.

frequency scopes, the heartbeat dynamics were the strongest (the absolute values of $K_{\tau(q)}$ are the largest), and also the most sensitive for distinguishing heart disease (discrepancies of parameter values between the two compared groups remained the greatest). When the heart rate of the mice decreases (e.g., GPD injected mice in Figure 4B), the above-mentioned maximal scale factor range moves higher, and meanwhile the sensitive frequency scope F_s of the ECG signal moves lower.

For each subject in Groups A, B, and C, we coarse-grained the V5-lead ECG time series of 300000 data points with a maximal scale factor 50, and then calculated the mass exponent spectrum curvature for each scale factor. The statistical averages in each group were then determined. The mean values with standard deviations of $K_{\tau(q)}$ for each of the three groups whose absolute magnitude were maximal among all the scale factors, the corresponding scale factor range and the exact position, and the relevant frequency scope of the ECG signal are shown in Table 1. For the negative value of $K_{\tau(q)}$, the absolute magnitude of it from the ECG of healthy mice was much larger than that of the other two groups, suggesting a greater bending degree of the $\tau(q) \sim q$ spectrum. The value of $K_{\tau(q)}$ from ECG of DOX injected mice was smaller than that of the other groups, with a spectrum tending to a straight line. These data suggest that the hearts of healthy mice were in optimal physiological states, and exhibited the strongest degree of nonlinear dynamics. By contrast, in the DOX injected mice whose heart muscles may be seriously damaged by drug toxicity, the complexity and fractal-like structure of the ECG attractors were even more reduced, which may be responsible for the marked loss of multifractality, and thus the linear ECG mass exponent spectrum. In GPD injected mice that may have exhibited less cardiac damage than DOX injected animals, the absolute value of $K_{\tau(q)}$ was larger than that of DOX injected mice, indicating a higher degree of heartbeat dynamics complexity.

The group means and group standard deviations of heart rates are shown in Table 1. The heart rate of DOX injected mice was much higher than that in the other two groups, while the heart rate for GPD injected mice decreased drastically to levels lower than healthy controls. Overall, these data suggest that (i) the heart rate of mice is not necessarily associated with the dynamic complexity of their hearts, but

only with the most sensitive frequency scope of ECG harmonic components in characterization of the strongest multifractal properties for this certain heartbeat condition. (ii) Only the cardiac health status is related to the nonlinear complexity of cardiac dynamics. Near the position of the extremum region of the nonlinear parameter, the complexity of the heart of healthy mice was higher than that of the pathological mice (i.e., positive correlation), while the complexity of healthy mice was lower than that of the pathological mice (negative correlation). For instance, in Figure 4A, the complexity becomes negatively correlated when the scale factor γ is larger than approximately 40. (iii) Around the extremum region of the $\tau(q) \sim q$ spectrum, this multifractal parameter is most susceptible to heart diseases. This parameter can efficiently discriminate between different physiological and pathological conditions, including the healthy and the heart diseased animals.

3.3 Statistical analysis of the classification results

We expressed the relevant classification results of our data as a 3×3 confusion matrix (eq. (8)), where the entries along the main diagonal were correct classifications, while entries other than those on the main diagonal were classification errors:

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}, \quad (8)$$

where c_{ij} denotes the number of samples classified into the j th group, but originally belonged to the i th group. The size of C is determined by the total sample number, i.e.,

$$N = \sum_{i=1}^3 \sum_{j=1}^3 c_{ij}, \text{ and restricted with constraint } c_i = \sum_{j=1}^3 c_{ij}, \\ c_i > 0, \quad c_{ij} \geq 0, \quad i = 1, 2, 3.$$

Usually, the two indexes of overall accuracy A and the Kappa coefficient are used to evaluate the confusion matrix, which are calculated by eqs. (9) and (10), respectively:

$$A = \frac{\sum_{i=1}^3 c_{ii}}{N} \times 100\%, \quad (9)$$

Table 1 Group averages and group standard deviations of $K_{\tau(q)}$ for the three classes of mice, the corresponding scale factor ranges, and the exact positions in which the absolute magnitude of the mean values achieves the maximum^{a)}

Groups	$K_{\tau(q)}$	Scale factor range	Scale factor position	Frequency scope F_s (Hz)	Heart rate (min)
A: DOX	-0.0018±0.0012	4–9	7	625–278	526.86±127.9
B: GPD	-0.0092±0.0008	10–15	10	250–167	419.90±89.59
C: Healthy	-0.0366±0.0016	7–12	9	357–208	434.64±97.25

a) Characteristic frequency scope F_s of ECG signals is calculated by the sampling frequency (5 kHz) divided by 2× scale factor range. The heart rates (mean±SD) in each group are also measured for a comparative study.

$$Kappa = \frac{N \sum_{i=1}^3 c_{ii} - \sum_{i=1}^3 c_{i+} c_{+i}}{N^2 - \sum_{i=1}^3 c_{i+} c_{+i}}, \quad (10)$$

where c_{ii} indicates correct classification numbers in each class. Symbol c_{i+} represents the sum of elements in row i , and c_{+i} is the corresponding value in column i .

Next, we performed a stratified hold-out sampling [30] to the data from Section 3.2 to obtain the confusion matrix and to quantify the classification rate (Table 2). First, we randomly removed 10% of the data sets from each of the three groups for testing, and used the remaining 90% for training to get the two critical values of K_{cr1} and K_{cr2} . Next, we classified each of the sequences removed for testing into one of the three groups according to these two values. The above steps were repeated 1000 times. The resulting confusion matrix is shown in Table 2. We were able to correctly classify 4392 out of 5000 subjects using this classifier, giving approximately 87.84% separation in the testing sets.

The results of the singularity spectrum width $\Delta\alpha$ [17,26] discussed in Section 1.2 are shown in Table 3. This classifier correctly classified 4196 out of 5000 subjects. The overall accuracy dropped slightly compared with the curvature $K_{\tau(q)}$, which reached 83.92%.

As well as the above classifiers based on the property of multifractality, we compared their performance with other classifiers in general [31] (Table 4). For the two parametric solutions of least squares (LS) and maximum likelihood (ML), the samples in the training set were assumed to obey

Table 2 The resulting three-class confusion matrix using $K_{\tau(q)}$ as the classification parameter^{a)}

Practical groups	Computed decisions		
	DOX (c_1)	GPD (c_2)	Healthy (c_3)
DOX (c_1)	1887	113	0
GPD (c_2)	291	1614	95
Healthy (c_3)	0	109	891

a) DOX injected mice were entirely separated from the healthy mice, and *vice versa*. There were 404 misclassifications between DOX injected mice and GPD injected mice, and 204 misclassifications between GPD injected mice and healthy mice.

Table 3 The three-class confusion matrix using $\Delta\alpha$ as the classification parameter^{a)}

Practical groups	Computed decisions		
	DOX (c_1)	GPD (c_2)	Healthy (c_3)
DOX (c_1)	1765	235	0
GPD (c_2)	346	1546	108
Healthy (c_3)	0	115	885

a) There were 581 misclassifications between DOX injected mice and GPD injected mice, and 223 misclassifications between GPD injected mice and healthy mice.

Table 4 Comparisons among the eight classifiers using different feature extraction methods^{a)}

Classification algorithms		Overall accuracy (%)	Kappa coefficient
Multifractality	$K_{\tau(q)}$	87.84	0.8098
	$\Delta\alpha$	83.92	0.7486
Parametric	Least squares (LS)	58.70	0.3499
	Maximum likelihood (ML)	64.13	0.4592
Non-parametric	AdaBoost	60.87	0.4107
	BP neural network	63.04	0.4562
	k -nearest neighbor (KNN)	75.00	0.6132
	Support vector machine (SVM)	61.96	0.3482

a) All algorithms were divided into either nonlinear or linear.

the law of normal distribution before estimation. For the non-parametric solutions, in the AdaBoost approach we use an iteration number of 100 times, and use the decision stumps as the weak learner type with a learner parameter of 100. In the back propagation (BP) neural network algorithm, five hidden units are used in the construction. The convergence criterion and the convergence rate were both 0.1. In the k -nearest neighbor (KNN) algorithm, the number of nearest neighbors was considered as $k=3$. In the support vector machine (SVM) approach, 'poly' kernel was chosen with a polynomial degree of two, and the solver type was selected as 'perceptron'. For all of these classifiers, the same holdout process presented above was used in the error estimation method, and one data number was drawn out randomly before calculations. Furthermore, all initial sequences were preprocessed with the feature selection method of principal component analysis (PCA) [32] to reduce their dimensions. The classification outcomes in this table suggest that classifiers based on nonlinear multifractal characteristics are all superior over other parametric and non-parametric classifiers for the analysis of our complex physiological time series.

4 Conclusion

In the present study, we examined the ECG mass exponent spectrum curvature $K_{\tau(q)}$ derived from healthy and drug-injected mice, as well as their distributions along with scale factors γ . The absolute magnitude of $K_{\tau(q)}$ from ECG of healthy mice was much larger than that for DOX and GPD injected mice, which exhibited some cardiac impairment due to drug toxicity. The absolute value from ECG of DOX injected mice is the least. As such, this parameter can effectively determine multifractal complexity in heartbeat dynamics, with larger absolute values indicative of better health condition, thus representing stronger multifractality and nonlinearity in heartbeat dynamics. The absolute values of $K_{\tau(q)}$ were different for each scale factor, and displayed

spatial distributions. Irrespective of the data length computed, the maximal value of $K_{\tau(q)}$ always existed within a certain scale factor range when the sampling frequency was fixed. In this range, the ECG signal exhibits the strongest nonlinear dynamic complexity, and is also most sensitive to indication of heart diseases. Then the sensitive characteristic frequency scope of the ECG signal can be determined based on relationship between the frequency and the scale factor. Our further investigations on data interpolations confirm that when the sampling frequency of the ECG signal is increased by n times, the above-mentioned maximal scale factor range would also increase by n times at the same time. However, the sensitive characteristic frequency scope of the ECG signal remains constant, and is also not changed with the data points calculated. Finally, the heart rate of mice was not associated with the nonlinear complexity of their cardiac dynamics, but only with the most sensitive frequency scope of the ECG signal in characterization of their heartbeat systems. With a decrease in heart rate, this frequency scope decreases to a lower position. In [21], we studied ECG signals from humans, whose heart rate is much lower than that of mice, and concluded that the sensitive characteristic frequency scope of them is in $F_s=(125-83)$ Hz with 1 kHz sampling frequency. This conclusion is just consistent with what we drew in this paper. Our findings are significant in determining the inherent laws in heartbeat dynamics and can also help in clinical applications.

This work was supported by the National Natural Science Foundation of China (Grant No. 61003169), the Ph.D. Programs Foundation of Ministry of Education of China (Grant No. 20090095120013), and the Technology Funding Project of China University of Mining and Technology (Grant No. 2008C004). We thank Dr. Wang Jun for helpful discussion.

- 1 Babloyantz A, Destexhe A. Is the normal heart a periodic oscillator? *Biol Cybern*, 1988, 58: 203–211
- 2 Bogaert C, Beckers F, Ramaekers D, et al. Analysis of heart rate variability with correlation dimension method in a normal population and in heart transplant patients. *Auton Neurosci*, 2001, 90: 142–147
- 3 Eckmann J P, Kamphorst S O, Ruelle D, et al. Liapunov exponents from time series. *Phys Rev A*, 1986, 34: 4971–4979
- 4 Richman J S, Moorman J R. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*, 2000, 278: H2039–H2049
- 5 Halsey T C, Jensen M H, Kadanoff L P, et al. Fractal measures and their singularities: The characterization of strange sets. *Phys Rev A*, 1986, 33: 1141–1151
- 6 Prasad R R, Meneveau C, Sreenivasan K R. Multifractal nature of the dissipation field of passive scalars in fully turbulent flows. *Phys Rev Lett*, 1988, 61: 74–77
- 7 Milde F, Romer R A, Schreiber M. Multifractal analysis of the metal-insulator transition in anisotropic systems. *Phys Rev B*, 1997, 55: 9463–9469
- 8 Ivanov P Ch, Amaral L A N, Goldberger A L, et al. Multifractality in human heartbeat dynamics. *Nature*, 1999, 399: 461–465
- 9 Stanley H E, Amaral L A N, Goldberger A L, et al. Statistical physics and physiology: Monofractal and multifractal approaches. *Physica A*, 1999, 270: 309–324
- 10 Grosse I, Herzel H, Buldyrev S V, et al. Species independence of mutual information in coding and noncoding DNA. *Phys Rev E*, 2000, 61: 5624–5629
- 11 Wang J, Ning X B, Chen Y. Multifractal analysis of electronic cardiogram taken from healthy and unhealthy adult subjects. *Physica A*, 2003, 323: 561–568
- 12 Ashkenazy Y, Havlin S, Ivanov P Ch, et al. Magnitude and sign scaling in power-law correlated time series. *Physica A*, 2003, 323: 19–41
- 13 Popivanov D, Jivkova S, Stomonyakov V, et al. Effect of independent component analysis on multifractality of EEG during visual-motor task. *Signal Process*, 2005, 85: 2112–2123
- 14 Kotani K, Struzik Z R, Takamasu K, et al. Model for complex heart rate dynamics in health and diseases. *Phys Rev E*, 2005, 72: 041904
- 15 Ma Q L, Ning X B, Wang J, et al. A new measure to characterize multifractality of sleep electroencephalogram. *Chin Sci Bull*, 2006, 51: 3059–3064
- 16 Ivanov P Ch, Ma Q L, Bartsch R P, et al. Levels of complexity in scale-invariant neural signals. *Phys Rev E*, 2009, 79: 041920
- 17 Amaral L A N, Ivanov P Ch, Aoyagi N, et al. Behavior-independent features of complex heartbeat dynamics. *Phys Rev Lett*, 2001, 86: 6026–6029
- 18 Costa M, Goldberger A L, Peng C K. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett*, 2002, 89: 068102
- 19 Costa M, Peng C K, Goldberger A L, et al. Multiscale entropy analysis of human gait dynamics. *Physica A*, 2003, 330: 53–60
- 20 Wang J, Ning X B, Ma Q L, et al. Multiscale multifractality analysis of a 12-lead electrocardiogram. *Phys Rev E*, 2005, 71: 062902
- 21 Yang X D, Ning X B, Wang J. Multifractal analysis of human synchronous 12-lead ECG signals using multiple scale factors. *Physica A*, 2007, 384: 413–422
- 22 Chhabra A, Jensen R V. Direct determination of the $f(\alpha)$ singularity spectrum. *Phys Rev Lett*, 1989, 62: 1327–1330
- 23 Ning X B, Bian C H, Wang J, et al. Research progresses in nonlinear analysis of heart electric activities. *Chin Sci Bull*, 2006, 51: 385–393
- 24 Grassberger P, Procaccia I. Characterization of strange attractors. *Phys Rev Lett*, 1983, 50: 346–349
- 25 Yang X D, He A J, Zhou Y, et al. Multifractal mass exponent spectrum of complex physiological time series. *Chin Sci Bull*, 2010, 55: 1996–2003
- 26 Ivanov P Ch, Amaral L A N, Goldberger A L, et al. From $1/f$ noise to multifractal cascades in heartbeat dynamics. *Chaos*, 2001, 11: 641–652
- 27 Wu S, Ko Y S, Teng M S, et al. Adriamycin-induced cardiomyocyte and endothelial cell apoptosis: *In vitro* and *in vivo* studies. *J Mol Cell Cardiol*, 2002, 34: 1595–1607
- 28 Outomuro D, Grana D R, Azzato F, et al. Adriamycin-induced myocardial toxicity: New solutions for an old problem? *Int J Cardiol*, 2007, 117: 6–15
- 29 Ning X B, Li D H, Ding S W, et al. Detection and identification of 12-lead synchronous ECG waveform (in Chinese). *J Nanjing Univ: Nat Sci Ed*, 2004, 40: 129–136
- 30 Witten I H, Frank E. Data mining: Practical machine learning tools and techniques. 2nd ed. New York: Morgan Kaufmann, 2005. 149–154
- 31 Bishop C M. Pattern Recognition and Machine Learning. New York: Springer, 2006
- 32 Turk M, Pentland A. Eigenfaces for recognition. *J Cognitive Neurosci*, 1991, 3: 71–86